Appl Intell (2008) 28: 195–205
DOI 10.1007/s10489-007-0062-2

Statistical properties analysis of real world tournament selection in genetic algorithms

S. Lee · S. Soak · K. Kim · H. Park · M. Jeon

Abstract Genetic algorithms (GAs) are probabilistic optimization methods based on the biological principle of natural evolution. One of the important operators in GAs is the selection strategy for obtaining better solutions. Specifically, finding a balance between the selection pressure and diversity is a critical issue in designing an efficient selection strategy. To this extent, the recently proposed real world tournament selection (*RWTS*) method has showed good performance in various benchmark problems. In this paper, we focus on analyzing characteristics of *RWTS* from the viewpoint of both the selection probabilities and stochastic sampling properties in order to provide a rational explanation for why *RWTS* provides improved performance. Statistical experimental results show that *RWTS* has a higher selection pressure with a relatively small loss of diversity and higher sampling accuracy than conventional tournament selection. The performance tests in a traveling salesman problem further confirm that the comparatively higher pressure and sampling accuracy, which are inherent in *RWTS*, can enhance the performance in the selection strategy.

Keywords Genetic algorithm · Real world tournament selection · Selection pressure · Sampling accuracy · Diversity

1 Introduction

Genetic algorithms (GAs) are stochastic techniques based on the mechanisms of natural selection and genetics; since the introduction of GAs by Holland [1], they have been widely used in various research areas. One of the most important issues in genetic algorithms is the tradeoff between *exploration* and *exploitation*. This tradeoff is reflected in three genetic operators: recombination, mutation, and selection. Generally, recombination and mutation explore the search space, whereas selection reduces the search area within the population by discarding poor solutions. However, selection is of primary importance because it is completely independent of the other GAs, and the fact that it affects both exploitation and exploration [17]. Therefore, it is also important to find a balance between exploration (i.e., poor solutions must have chance to go to the next generation) and exploitation (i.e., good solutions go to the next generation more frequently than poor solutions) within the mechanism of the selection strategy. In GAs, there are various selection methods, such as proportionate selection [1], linear ranking selection (*LRS*) [2], exponential ranking selection (*ERS*) [3], and tournament selection (*TS*) [4]. Among these selection schemes, *TS* is the most popular in genetic algorithms due to its efficiency and simple implementation.

S. Lee · K. Kim · M. Jeon (✉)
School of Information and Mechatronics, Gwangju Institute of Science and Technology, 1 Oryong-dong, Buk-gu, Gwangju 500-712, South Korea
e-mail: mgjeon@gist.ac.kr

S. Lee
e-mail: yashin96@gist.ac.kr

K. Kim
e-mail: mkkim@gist.ac.kr

S. Soak
Information Examination Team, KIPO, Gov. Complex Daejeon Bldg. 4, 920 Dunsandong, Seo-gu, Daejeon, South Korea
e-mail: soakbong@hotmail.com

H. Park
College of Computing, Georgia Tech., Atlanta, GA 30332, USA
e-mail: hpark@gatech.edu

Recently, Soak et al. [5–7] introduced a real world tournament selection (*RWTS*) scheme that can mimic an actual sporting tournament. Conventional *TS* selects only one pair of chromosomes from a population, and only the winner goes into the mating pool, while whole population is maintained in the competition pool. However, *RWTS* pairs all chromosomes in the population, and then only the winners of each competition go into the mating and competition pools for the next level (see Fig. 1). In [5–8], *RWTS* was tested for various benchmark problems such as tree-based optimization, function optimization, multidimensional knapsack, and traveling salesman problems. Experimental results showed that *RWTS* provides a superior performance compared to the conventional *TS* in all test instances, except when using edge set encoding in tree-based optimization problems, with respect to the solution quality, standard deviation, and computational time. However, these studies did not provide any detailed mathematical analysis of *RWTS* [5–7], and Lee et al. [8] gave only probability distribution of *RWTS*.

The selection process can be analyzed as a two-step procedure consisting of assigning selection probabilities and applying a sampling strategy. The analysis of selection probabilities is well described in Julstrom [9] and Bäck [10]. For instance, Julstrom has suggested a selection probability according to each individual's rank for two kinds of rank-based selections and four kinds of tournament selections, and also shown the equivalence between rank-based selection and tournament selection at specific parameter settings. Moreover, we have already derived a numerical formula that suggests that the selection probability of *RWTS* can be calculated according to an individual's rank [8].

An extensive survey regarding sampling strategies such as roulette wheel sampling (*RW*) and stochastic universal sampling (*SUS*) can be found in [11]. Here, the authors state that a smaller sampling error is preferable for reducing genetic drift, where the genetic drift causes the population to be trapped in a local optimum. The effects of genetic drift are well described in [12, 13]. Recently, Schell et al. [14] proposed an χ^2 measure to analyze the stochastic properties of sampling algorithms; the χ^2 Goodness-of-Fit measure can provide a calibrated scale of accuracy, ranging from almost deterministic to fully stochastic sampling algorithms. They further state that sampling which yields a small sampling error has high *sampling accuracy*.

Selection pressure is one of the most important characteristics of the selection strategy, since it is deeply concerned with the balance between exploitation and exploration [16]. It is defined as the ratio of the best individual's selection probability to the average selection probability of all individuals in the selection pool [17]. In [8], it is shown that *RWTS* has a higher selection pressure than a conventional *TS* with a tournament size of 2. However, the selection pressure of *TS* can be raised by increasing the tournament size [16].

In this paper, we focus on the analysis of characteristics of *RWTS* and *TS* from the viewpoint of both the selection probabilities and stochastic sampling properties to investigate their selection pressure and sampling accuracy. In *TS*, we will examine two cases: the general case (tournament size of 2), and the special case. The special case is the situation where the selection pressure of *TS* is similar to that of *RWTS*, achieved by adjusting the tournament size. Furthermore, we will test selection strategies on a traveling salesman problem to investigate how these two characteristics affect the performance and the diversity of GAs.

This paper is organized as follows. The two-step and tournament selections are overviewed in Sect. 2, and *RWTS* is described in Sect. 3. Then, an empirical analysis and conclusions are given in Sect. 4 and Sect. 5, respectively.

2 Two-step selection and tournament selection

The selection process in genetic algorithms can be divided into two steps. The first step is to assign a selection probability to each individual with regards to its fitness; examples are given in Sect. 2.1. The expected number e_i of the offspring rate of each individual i for the next generation can easily be calculated by multiplying the individual's selection probability p_i into N as $e_i = N \cdot p_i$, where N is the population size.

The second step is sampling, which entails selecting N individuals from the current population through the use of a sampling algorithm. In this process, the selection probabilities are reflected in such a way that o_i, the number of offspring of individual i, has the expectation e_i, $E[o_i] = e_i$. The absolute differences between o_i and e_i, $|o_i - e_i|$, will vary according to the stochastic properties of the sampling process. This difference is referred to as the sampling error.

Generally, there are two kinds of selection methods, the separate two-step method mentioned above, and a combined two-step method, such as *TS*. In the combined case, selection probabilities and sampling properties are intrinsic to the selection scheme. We will briefly review the separate two-step method and *TS* in the remainder of this section.

2.1 Assignment of probability

Various schemes for assigning selection probabilities to individuals according to their fitness have been discussed in Bäck [10] and Blickle et al. [15]. In this section, however, only rank-based algorithms are reviewed.

Linear ranking selection (*LRS*) was first introduced by Baker [2]. In this selection process, we let the N individuals in a population be sorted according to their increasing fitness, from the worst individual ranked 1 to the best ranked N, where the best individual is assigned a value s, $1 < s < 2$, and the worst individual is assigned $(2-s)$. The remaining individuals are assigned values linearly proportional to their ranks, with a slope of $\frac{2(s-1)}{N-1}$. Then, the selection probability of *LRS* for the ith ranked individual can be given by

$$p_i^{LRS} = \frac{1}{N}\left((2-s) + \frac{2(i-1)(s-1)}{N-1}\right). \tag{1}$$

Exponential ranking selection (*ERS*) is identical to *LRS*, except for the non-linear assignment of probabilities to individuals. In this case, the rank-based values that differ by a constant ratio r are assigned to the individuals, such that $0 < r < 1.0$ ($r \simeq 1.0$), where the best individual is assigned a value of 1 and the ith ranked individual is assigned the value r^{n-i}. Then, the selection probability of *ERS* for the ith ranked individual can be given by

$$p_i^{ERS} = \frac{r^{N-i}(1-r)}{1-r^N}. \tag{2}$$

2.2 Sampling

A sampling algorithm is required in a two-step selection scheme such as rank-based selection in order to select individuals for recombination. It fills the mating pool with copies of the current population's individuals; the copy process reflects the selection probabilities p_i, and the expectation of each offspring o_i is $E[o_i] = e_i$. Then, in this paper, two popular methods, *RW* and *SUS*, are used for testing.

The mechanism of *RW* introduced by Holland [1] is as follows. On the roulette wheel, each solution is assigned to a slice with a size proportional to its value of probability p_i. A single marker is placed at the edge of the biased roulette wheel, and the roulette wheel is spun N times to allow the marker to successively select individuals. However, the *RW* sampling has the potential to generate large sampling errors, in the sense that the final number of offspring allocated to a string might significantly vary from the expected number.

Thus, in order to reduce the sampling error, Baker introduced *SUS* [2]. *SUS* is almost the same as *RW*, except for the number of markers. Instead of the single marker used in *RW*, *SUS* makes use of N markers spaced evenly around the edge of the roulette wheel. The roulette wheel is spun only once, and all individuals indicated by the N markers are inserted into the mating pool.

2.3 The χ^2 goodness-of-fit measure

The χ^2 measure was first introduced by Schell et al. [14] as a tool for measuring the average accuracy; the average difference between the expectation of each offspring and the actual numbers of each offspring obtained by the sampling algorithms. The stochastic properties of a sampling method are described by the sample mean, sample variance, and empirical distribution function of the realizations. The details of this method are as follows.

First, the individuals are grouped into disjoint c classes $\{\mathcal{C}_1, \ldots, \mathcal{C}_c\}$, where $\mathcal{C}_j \subset \{1, \ldots, N\}$ and $\bigcup_{j=1}^{c} \mathcal{C}_j = \{1, \ldots, N\}$; let $\mathcal{E}_j = \sum_{i \in \mathcal{C}_j} e_i$ and $\mathcal{O}_j = \sum_{i \in \mathcal{C}_j} o_i$ denote the overall expectation and the overall observation, respectively, which represents the overall number of actual offsprings, after the sampling process, in each class. Ideally, $\mathcal{E}_j$ should be of the order N/c for $1 \le j \le c$, so that each class contributes the same average number of offspring. Furthermore, $\mathcal{E}_j \ge 10$ for $1 \le j \le c$ is required to obtain the desired stochastic accuracy. Schell et al. defined the measure for determining the accuracy of the sampling algorithms as the chi-square test as

$$\chi := \sum_{j=1}^{c} \frac{(\mathcal{E}_j - \mathcal{O}_j)^2}{\mathcal{E}_j}. \tag{3}$$

In the *SUS* case, $\chi \approx 0$ because of *SUS*'s characteristic of minimizing the differences $e_i - o_i$. In the *RW* case with the aforementioned constraints (i.e., $\mathcal{E}_j \ge 10$), however, χ should be approximately chi-square distributed with $c-1$ degrees of freedom, since this is the asymptotic distribution of χ under multinomial distributed o_i when $N \to \infty$. Specifically, on average, χ should be close to its expectation: $\chi \approx E[\chi^2_{c-1}] = c-1$. If the population size is small, the average value is slightly higher because of the crude approximation of $\mathcal{E}_j$. However, this effect vanishes if $N \ge 250$ [14]. Note that a higher χ implies a lower accuracy.

2.4 Tournament selection

Tournament selection is an integral method, where the two-step selection process is combined into one step, as previously mentioned. Instead of assigning selection probability to each individual, the individuals are randomly copied into a tournament-list, with size $t_s \ge 2$. The fittest individual in the tournament-list is then inserted into the mating pool. This procedure is repeated until the mating pool is full.

In the following subsections, four kinds of *TS* will be briefly reviewed.

2.4.1 Tournament selection with replacement (TSR)

In TSR, an individual is randomly sampled from the population, and its copy is inserted into the tournament-list, while the original individual is kept in the population. This sampling process is repeated t_s times, and finally, from among the t_s individuals in the tournament-list, the fittest individual is inserted into the mating pool.

In order to calculate the selection probabilities of *TSR*, it is assumed that all individuals have different fitness, and that the population is ordered in such a way that the lowest rank 1 is assigned to the worst individual, and the highest rank N is assigned to the best of the N individuals. The probability that the ranks of all sampled t_s contestants are lower than or equal to i is $(\frac{i}{N})^{t_s}$, due to the fact that the contestants are sampled with replacement, that is, the samplings are independent. As a minimum, this implies that more than one ith ranked individual must be included in the tournament group. So, the selection probability of the ith ranked individual in *TSR* is the difference between the probability that the ranks of all contestants are lower than or equal to i, and the probability that the ranks of all contestants are lower than or equal to $i-1$. Thus, the selection probability of *TSR* for the ith ranked individual can be given by

$$p_i^{TS_R} = \frac{i^{t_s} - (i-1)^{t_s}}{N^{t_s}}. \tag{4}$$

Julstrom [9] has shown that distribution $p_i^{TS_R}$ is equivalent to p_i^{ERS} for certain values of t_s and corresponding N, r.

2.4.2 *Tournament selection with partial replacement (TSpR)*

In *TSpR*, the population is copied t_s times, and each copy is randomly permuted. The copies are then concatenated to a tournament-list of size $t_s \cdot N$. Then, for each tournament, the t_s elements are successively removed from the tournament-list, and the winner is inserted into the mating pool.

Let

$$\binom{a}{b} = \frac{a!}{b!(a-b)!}$$

be the number of possibilities for selecting b elements out of a. The ith ranked individual has

$$\frac{\binom{i-1}{t_s-1}}{\binom{N-1}{t_s-1}}$$

chances of being selected in a single tournament if it is contained in this tournament. Thus, the average selection probability of that individual can be given by

$$p_i^{TS_{PR}} = \frac{\binom{i-1}{t_s-1}}{\binom{N}{t_s}}. \tag{5}$$

2.4.3 *Tournament selection without replacement (TSoR)*

The selection process of *TSoR* is almost the same as *TSR*, except that there is no duplication of individuals in the tournament-list. Then, as the number of all combinations of the tournament groups is $\binom{N}{t_s}$, the selection probability of *TSoR* for the ith ranked individual can be given by

$$p_i^{TS_{oR}} = \frac{\binom{i-1}{t_s-1}}{\binom{N}{t_s}}. \tag{6}$$

2.4.4 *Probabilistic 2-tournament selection (pTS)*

pTS was presented by Julstrom [9]. In this case, from a population, two individuals without replacement are chosen at random, and the tournament winner will be inserted into the mating pool with a probability q, such that $0.5 < q < 1.0$. Then, with the probability $(1-p)$, the tournament loser will be inserted into the mating pool. Thus, the selection probability of *pTS* for the ith ranked individual can be given by

$$p_i^{TS_P} = \frac{2(i-1)}{N(N-1)}q + \frac{2(N-i)}{N(N-1)}(1-q). \tag{7}$$

Julstrom further posited that the distribution $p_i^{TS_P}$ is equivalent to p_i^{LRS} for certain values $0.5 < q < 1.0$, and their corresponding s.

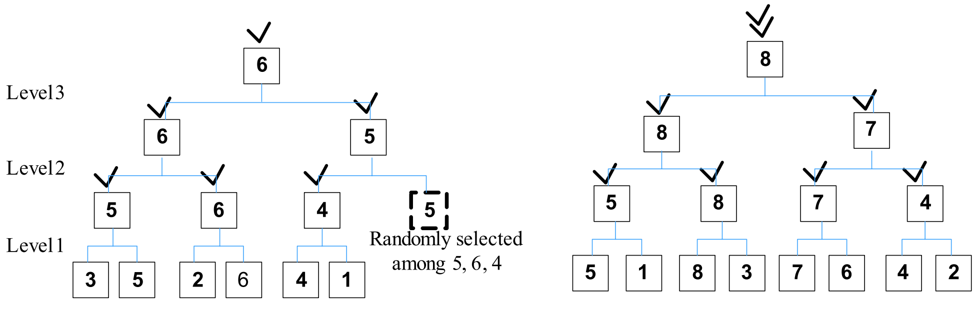
3 Real world tournament selection

To select a champion from among many players in a sports game, a tournament method is widely used. If one wins a competition, then one survives; if not, one is eliminated. *RWTS* is derived from such a real world tournament competition.

In *RWTS*, each individual in the population is sequentially paired with a neighbor without replacement or omission. If the last individual has no neighbor (i.e., the population is *odd*), this individual competes with an individual randomly selected from among the same tournament level. When all competitions in the present tournament level are completed, only the winners are inserted into the mating pool and go on to the next tournament level. The process is repeated until the mating pool is full. Figure 1 shows how *RWTS* works with 6 individuals and 8 individuals, respectively. Note that the number inside the box indicates the individual's fitness value, and a dashed line box indicates an individual selected randomly from the same tournament level to mate with the individual with no neighbor. As the result of the tournament competition, all checked individuals (i.e., the total is the population size) are inserted into the mating pool.

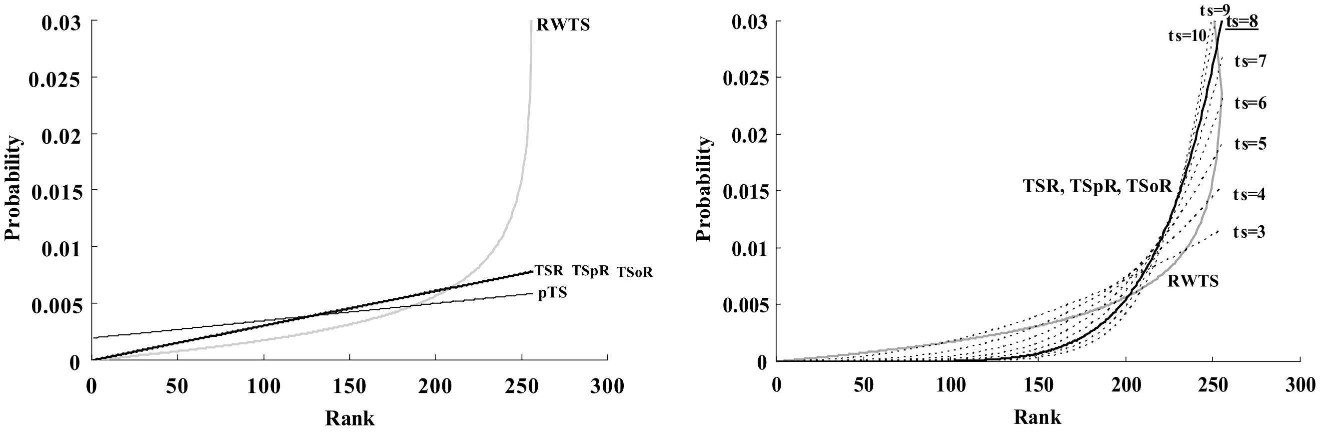
In [5–8], *RWTS* was applied to a minimum spanning tree, tree-based combinatorial optimization, function opti-

Fig. 1 Description of *RWTS*



(b) 6 individuals

(d) 8 individuals



(a) General parameter set ($t_s = 2$)

(b) Special parameter set : fitted on RWTS

Fig. 2 Comparison of the distributions of the selection probabilities

mization, traveling salesman, and multidimensional knapsack problems. The experimental results showed that *RWTS* is superior to the conventional *TS* with regards to solution quality, computational time, and stability.

The selection probability of the ith ranked individual has been previously well described [8]. Here, we give the final formula as follows

$$p_i^{rwts} = \frac{1}{N} \sum_{k=1}^{\lfloor \log_2 i \rfloor} \left(\prod_{l=1}^{2^k - 1} \frac{i-l}{N-l} \right), \quad \text{for } i = 1, \ldots, N-1$$
$$= \frac{\lfloor \log_2 N \rfloor + 1}{N}, \quad \text{for } i = N, \qquad (8)$$

where $\lfloor x \rfloor$ is the floor of the number x. It is assumed that the population size is $N = 2^n$, as in Fig. 1(b), and that all individuals have a different fitness; the population is ordered in such a way that the worst individual is assigned the lowest rank 1, and the best individual is assigned the highest rank N. For $N = 2^n$, *RWTS* performs tournament competitions on n levels.

4 Empirical analysis

In this empirical analysis, we will show the probability distribution and stochastic properties of *RWTS*, and these two characteristics are then compared with those of *TS*. Next, we will present the performance and diversity tests of *RWTS* and *TS* in a traveling salesman problem (TSP). Here, for our computational experiments, we used a Pentium IV with a 2.4 GHZ CPU and 1 GByte of memory, and Visual C++ as the program language.

4.1 Probability distribution comparison between *RWTS* and *TS*

In this subsection, we investigate the difference between *RWTS* and *TS* by comparing the selection probability distribution with two tournament sizes, the general case ($t_s = 2$) and a special case ($t_s = 8$), in which the selection pressure of *TS* is similar to that of *RWTS*. The purpose of the special case is to observe the difference between the selection probability distribution of *RWTS* and *TS* under similar selection pressure conditions. Figure 2 shows the distribution of the selection probabilities of *RWTS* and *TS* with *pop_size*

$(N) = 256$. This distribution was obtained by plotting the ith probability of the selection strategies (p_i^{SS}) from $i = 1$ to $i = N$, where the superscript SS is the set of $rwts$, TS_R, TS_{pR}, TS_{oR}, and TS_p. In the graph, the value on the vertical axis indicates the probability that the ranked individual on the horizontal axis is chosen for the matting pool. Note that the lowest rank 1 indicates the worst individual, and the highest rank N indicates the best individual.

Figure 2(a) shows the distribution of the selection probabilities of *RWTS* and four types of *TS* (*TSR*, *TSpR*, *TSoR*, and *pTS*) with the tournament size $t_s = 2$. In the case of *pTS*, we used $q = 0.75$, which is the middle value between 0.5 and 1.0. From the graph, we can see that below the 85% rank, the probabilities of *RWTS* are relatively similar (smaller) with those of *TS*; through above the 85% rank, *RWTS* has significantly higher selection probabilities than *TS*. This higher value implies that *RWTS* has a *higher selection pressure* than *TS* at $t_s = 2$.

In order to create a condition of similar selection pressure between *RWTS* and *TS*, the selection pressure of *TS* is adjusted by increasing the tournament size t_s. Figure 2(b) illustrates the adjusted selection probability distributions from $t_s = 3$ to $t_s = 10$. Note that the gray line indicates the selection probability distribution of *RWTS*, and the black and dashed lines indicate those of *TS* with $t_s = 8$ and the others, respectively. Theoretically, it is impossible to fit the distribution of *pTS* to that of *RWTS*. From the graph, we have selected $t_s = 8$ is the best choice for making a special case. In this case, we can see that *RWTS* has higher probabilities below the 80% rank than those of *TS*; specifically, the probability of *TS* is zero below the 50%. This means that *TS* with $t_s = 8$ has no chance to survive in the population pool at below 50% rank solutions. From these facts, we can predict that *RWTS* has a greater potential to provide a *higher diversity* to populations than *TS* at similar high selection pressures. A more detailed explanation for diversity will be described in Sect. 4.4.

4.2 Stochastic properties comparison between *RWTS* and *TS*

In these numerical experiments, we examine the stochastic properties of *RWTS* using the χ measure, and then compare the results with those of conventional *TS*. Here, we experimented with several different kinds of random number generators, and various parameter combinations of population size, number of classes, and number of tests under the constraints previously mentioned in Sect. 2.3. However, a comparison of all tests shows only a slight difference in results; there is no significant difference among the results of different kinds of random number generators and among the results of a different number of classes. The results of the analysis of a large population size displays a very slightly smaller sampling error than in a small population, but the difference is negligible. In addition, the results of large number of tests gives a smoother *EDF* line, but there is no difference in sampling error. In this paper, therefore, we present the results with only one parameter setting in order to compare its results to that of conventional *TS* [14]; the random number generator = *drand*48, the population size $N = 256$, the number of classes $c = 10$, and the number of tests $tn = 256$.

Table 1 presents the index sets used for the probability distributions of *RWTS* and *TS*s. The corresponding overall expectations that are very close to 256/10. We denote the results of χ measured by $\chi^{(D,S)}$, where D is a type of *TS* in p_i^D in $\{TS_R, TS_{pR}, TS_P, rwts\}$, and S is a type of sampling algorithm in $\{RW, SUS\}$. For the integral methods *TSR*, *TSpR*, *TSoR*, *pTS*, and *RWTS*, we denote the corresponding statistics as $\chi^{(TSR_{t_s})}$, $\chi^{(TSpR_{t_s})}$, $\chi^{(TSoR_{t_s})}$, $\chi^{(pTS)}$, and $\chi^{(RWTS)}$, where subscript t_s is the tournament size.

The main objective of the following numerical test is to estimate the expectation and variance of $\chi^{(D,S)}$ from a sequence $(\chi_k^{(D,S)})_{1 \le k \le tn}$ such that the mean accuracy and the spread of the sampling algorithm S are given in terms of the expectation and the variance of χ. For each test, the population is generated with both fixed and pairwise different fitness values, and the selection probabilities are assigned according to probability distribution D. Then, the sampling method S is applied to obtain instances of o_i, $\mathcal{O}_j$, and $\chi^{(D,S)}$, respectively.

From the sequence $(\chi_k^{(D,S)})_{1 \le k \le tn}$, the sample mean and the sample standard deviation can be calculated as

$$\hat{e}^{(D,S)} = \frac{1}{tn} \sum_{k=1}^{tn} \chi_k^{(D,S)},$$

$$\hat{\sigma}^{(D,S)} = \sqrt{\frac{1}{tn-1} \sum_{k=1}^{tn} (\chi_k^{(D,S)} - \hat{e}^{(D,S)})^2}.$$

Here, $\hat{e}^{(D,S)}$ and $\hat{\sigma}^{(D,S)}$ provide good estimates of the accuracy and spread in terms of the expectation $E[\chi^{(D,S)}]$ and the standard deviation $\sqrt{\mathrm{Var}[\chi^{(D,S)}]}$, respectively. Finally, the empirical distribution function (*EDF*) can be compared to the theoretical chi-square distribution χ^2_{c-1} of *RW* by means of the two-sided Kolmogorov–Smirnov (KS) test

$$EDF^{(D,S)}(t) = \frac{1}{tn} \cdot |\{\chi_k^{(D,S)} \le t,\ 1 \le k \le tn\}|,$$
$$t \in [0, \infty)$$

where $|\{\cdot\}|$ refers to the *cardinality of* $\{\cdot\}$.

Figure 3 shows the behaviors of $EDF^{(RWTS)}$ (dark thick line), $EDF^{(rwts,RW)}$ (gray thick line), and $EDF^{(rwts,SUS)}$ (dashed thick line), which almost overlap with the left axis and the upper part of the bounding box in the graph, for

Table 1 Classes $\mathcal{C}_j$ and overall expectations $\mathcal{E}_j$ for *RWTS*

j	*RWTS*		*TSR*$_2$		*TSR*$_8$		*TSpR*$_2$		*TSpR*$_8$		*TSoR*$_2$		*TSoR*$_8$		*pTS*	
	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$	$\mathcal{C}_j$	$\mathcal{E}_j$
1	1–110	25.66	1–81	25.63	1–192	25.63	1–81	25.41	1–193	25.76	1–81	25.41	1–193	25.76	1–44	25.71
2	111–150	26.00	82–114	25.14	193–209	24.90	82–114	25.11	194–210	25.46	82–114	25.11	194–210	25.46	45–79	25.87
3	151–77	25.52	115–140	25.80	210–220	25.63	115–140	25.80	211–221	26.36	115–140	25.80	211–221	26.36	80–109	26.00
4	178–198	26.24	141–162	25.95	221–228	25.19	141–162	25.97	222–229	26.00	141–162	25.97	222–229	26.00	110–135	25.39
5	199–214	25.37	163–181	25.46	229–235	27.74	163–181	25.48	230–235	24.22	163–181	25.48	230–235	24.22	136–159	25.79
6	215–227	25.63	182–198	25.17	236–240	23.68	182–198	25.20	236–240	23.83	182–198	25.20	236–240	23.83	160–181	25.62
7	228–237	24.35	199–214	25.75	241–245	27.40	199–214	25.79	241–245	27.63	199–214	25.79	241–245	27.63	182–202	26.23
8	238–245	24.18	215–229	25.96	246–249	24.92	215–229	26.00	246–249	25.18	215–229	26.00	246–249	25.18	203–221	25.22
9	246–252	27.70	230–243	25.82	250–253	27.89	230–243	25.86	250–253	28.22	230–243	25.86	250–253	28.22	222–239	25.20
10	253–256	25.35	244–256	25.34	254–256	23.04	244–256	25.39	254–256	23.35	244–256	25.39	254–256	23.35	240–256	24.97

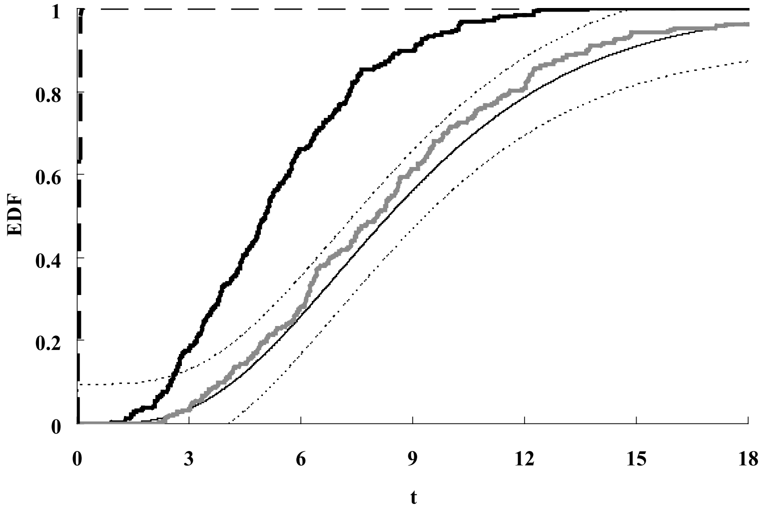


Fig. 3 Comparison of *RWTS*, with *RW* and *SUS* sampling based on probabilities

a population size $N = 256$ and number of tests $tn = 256$. These schemes are compared with the theoretical χ^2_{c-1} distribution (dark thin line) and the 99 % confidence band under the hypothesis of *RW* (dashed thin double line). Here, the range for the values t is $[0, 18]$, so the expectation of $\chi^{(D,RW)}$ is $c - 1$ $(= 9)$. The corresponding estimates of $\hat{e}$, $\hat{\sigma}$, and the KS statistic are given in Table 2.

The confidence band has been derived from the KS distribution in [18]. Note that if $EDF^{(D,S)}$ does not lie completely within this band, the sampling algorithm S is not the same as the *RW* sampling with a probability that is greater than or equal to 99 %. This statistical statement can be confirmed by the KS statistic value, which has to be smaller than about 1.63 for the same level of significance, provided that

Fig. 4 Comparison of sampling between *RWTS* and *TS*

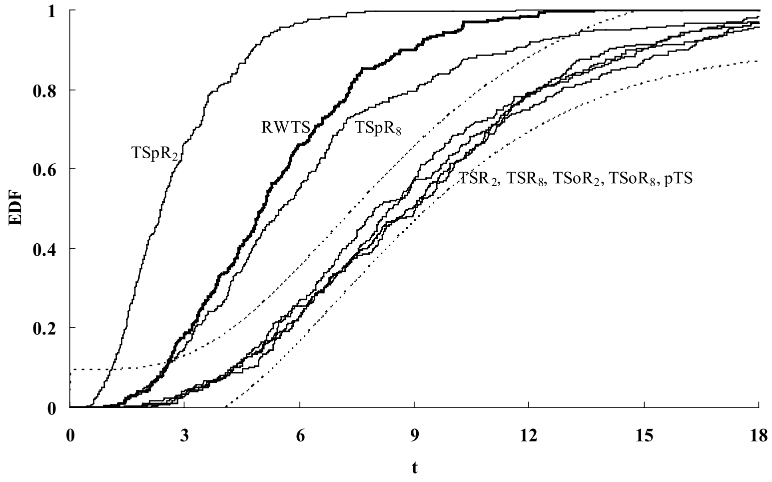


Table 2 Sample mean and standard deviations for *EDF* in Fig. 3

	$RWTS$	$(rwts, RW)$	$(rwts, SUS)$
$\hat{e}$	5.359	8.529	0.079
$\hat{\sigma}$	2.462	4.052	0.026
KS	**6.791**	1.123	**16.0**

$tn \geq 40$. In our work, if the test rejects the *RW* hypothesis, the KS value is printed in boldface.

Therefore, from Table 2, we can confirm that *RWTS* differs from *RW*; the relatively high KS value of *RWTS* illustrates the fact that *RWTS* has a higher accuracy and smaller spread than the *RW* sampling. However, *RWTS* also has lower accuracy than *SUS*. Thus, a more symbolic representation of a comparison of accuracy from among the three methods is: $(rwts, RW) < RWTS < (rwts, SUS)$. The results of $EDF^{(rwts,RW)}$ and $EDF^{(rwts,SUS)}$ are almost the same as [14] in the case of (D, RW) and (D, SUS).

Figure 4 shows a comparison of the behaviors of $EDF^{(RWTS)}$ and $EDF^{(TS)}$. Note that TSR_2, TSR_8, $TSoR_2$, $TSoR_8$, and *pTS* follow the *RW* sampling hypothesis stated above. The corresponding estimates of $\hat{e}$ and $\hat{\sigma}$, and the KS statistic value are given in Table 3. From this table, we can see that though *RWTS* has a higher sampling accuracy than TSR_2, TSR_8, $TSoR_2$, $TSoR_8$, $TSpR_8$, and *pTS*, it has a lower sampling accuracy than $TSpR_2$. Hence, a more symbolic description of a comparison of the sampling accuracies from among the four methods is: $pTS \approx TSR_2 \approx TSR_8 \approx TSoR_2 \approx TSoR_8 < TSpR_8 < RWTS < TSpR_2$.

In [11], Baker stated that a small sampling error is desirable for reducing genetic drift, where genetic drift causes solution to converge on a unintended point due to the accumulation of sampling errors. In other words, the difference between the actual sampling frequency and the calculated selection probabilities will eventually cause GA to converge to a local optimum. Based on this statement, *RWTS* has better properties than *TS* with respect to sampling accuracy, except for $TSpR_2$.

4.3 Performance comparison between *RWTS* and *TS*

Through the above two subsections, we can determine the relative characteristics of *RWTS* and *TS* with respect to the selection pressure and sampling accuracy. Now, we test the performance of *RWTS* by applying it to TSP, and then comparing its performance to the four *TS* with $t_s = 2$ and $t_s = 8$, with a population size of 256. The benchmark problem of TSP is well described in TSPLIB [19]. Here, we have selected the five sample TSP problems, eil51, eil76, eil101, gr202, and pa561 for the performance tests.

The parameters of a GA test are as follows:

(1) encoding: random key representation [20]
(2) crossover: uniform crossover [21] with 0.6 probability
(3) mutation: swap mutation with 0.1 probability
(4) population size: 256
(5) number of generations: 400

For the encoding scheme, we use a system of random keys (RK) which is identical to the one proposed by Bean [20]. An important feature of RK is that all offspring formed by genetic operations (i.e., crossover and mutation) are feasible solutions. This is accomplished by moving much of the feasibility issue into the objective function evaluation; if any random key vector can be interpreted as a feasible solution, then any subsequent crossover and mutation vector is feasible. Through the dynamics of the genetic algorithm, the system learns the relationships between the random key vectors and solutions with appropriate objective function values.

As genetic operators, we use a uniform crossover and a swap mutation. In the uniform crossover scheme, individual genes of two parents in the string are compared, and then the genes are swapped with a fixed probability. In the swap mutation, two random points in the string are selected, and

Table 3 Comparison of *RWTS* with *TS* stochastic properties based on sample mean and standard deviation

	RWTS	*TS*						
		TSR_2	TSR_8	$TSpR_2$	$TSpR_8$	$TSoR_2$	$TSoR_8$	*pTS*
$\hat{e}$	5.359	8.802	9.260	3.260	7.352	9.321	8.938	8.916
$\hat{\sigma}$	2.462	4.297	4.178	1.716	8.384	4.435	4.169	4.066
KS	**6.791**	0.929	1.257	**10.920**	**5.420**	0.781	0.641	0.476

Table 4 Experimental results of TSP (population size = 256, number of generations = 400)

Sel.	eil51					eil76					eil101				
	Min.	Ave.	Max.	STD.	CPU	Min.	Ave.	Max.	STD.	CPU	Min.	Ave.	Max.	STD.	CPU
RWTS	0.63	1.00	1.45	0.19	1.85	0.77	1.00	1.17	0.09	2.81	0.90	1.00	1.12	0.06	3.85
TSR_2	0.81	1.17	1.51	0.17	1.88	0.92	1.11	1.30	0.07	2.89	0.99	1.08	1.16	0.04	3.99
TSR_8	0.83	1.12	1.45	0.17	1.86	0.89	1.06	1.22	0.10	2.83	0.95	1.05	1.18	0.05	3.86
$TSoR_2$	0.82	1.13	1.44	0.17	1.86	1.06	1.17	1.32	0.07	2.91	0.97	1.11	1.23	0.06	3.98
$TSoR_8$	0.67	1.06	1.52	0.18	1.88	0.77	1.03	1.23	0.11	2.85	0.94	1.01	1.09	0.04	3.92
$TSpR_2$	0.64	0.98	1.20	0.13	1.95	0.87	1.01	1.15	0.07	3.04	0.90	1.00	1.09	0.04	4.14
$TSpR_8$	0.64	1.02	1.44	0.18	2.24	0.80	1.03	1.17	0.08	3.23	0.90	0.98	1.07	0.04	4.29
pTS	1.04	1.35	1.71	0.15	1.86	1.17	1.35	1.50	0.08	2.87	1.13	1.23	1.32	0.05	3.94

Sel.	gr202					pa561				
	Min.	Ave.	Max.	STD.	CPU	Min.	Ave.	Max.	STD.	CPU
RWTS	0.95	1.00	1.03	0.02	8.11	0.99	1.00	1.01	0.01	23.27
TSR_2	1.04	1.07	1.11	0.02	8.28	1.03	1.04	1.05	0.00	23.83
TSR_8	0.98	1.03	1.08	0.02	8.18	0.98	1.00	1.01	0.01	23.75
$TSoR_2$	1.00	1.08	1.12	0.03	8.32	1.03	1.04	1.04	0.00	23.90
$TSoR_8$	0.95	1.00	1.03	0.02	8.20	0.98	1.00	1.00	0.01	23.89
$TSpR_2$	0.91	0.96	1.01	0.02	8.48	0.95	0.96	0.97	0.01	24.20
$TSpR_8$	0.93	1.00	1.04	0.03	8.61	0.98	0.99	1.00	0.00	24.61
pTS	1.12	1.18	1.22	0.02	8.10	1.06	1.08	1.09	0.01	23.41

then the two genes at these positions are interchanged, where the individual for mutation is selected with a fixed probability. In this test, we use 0.6 and 0.1 as the probabilities for crossover and mutation, respectively.

All tests were repeated thirty times. Table 4 shows the results of the performance tests, where Min. indicates the minimum value of the best values from among the 30 trials; Ave. indicates the average value of the best values of the 30 trials; Max. indicates the maximum value of the best values from among the 30 trials; STD. indicates the standard deviation of the 30 trials and CPU indicates the computational time (sec) for 256 × 400 evaluations. Here, all values of Min., Ave., and Max. are normalized based on the *RWTS* Ave. value. Since TSP is a minimization problem, smaller values of Ave. indicate better selection methods from among the eight selection methods.

From the above results, we can confirm the fact that *RWTS* and *TSpR* outperform the others. Specifically, we can see that the case of $t_s = 8$ outperforms $t_s = 2$ in *TS*, except for *TSpR*. In *TSpR*, one possible reason why $t_s = 2$ outperforms $t_s = 8$ is that the sampling accuracy with $t_s = 2$ is higher than that of $t_s = 8$. These results indicate that a comparatively high pressure and high sampling accuracy assist in improving the performance in the selection method. In literature, it was also mentioned that high pressure [17] and high sampling accuracy [14] are desirable. However, keep in mind that the choice of high selection pressure causes a corresponding loss of diversity.

4.4 Diversity comparison between *RWTS* and *TS*

In Sect. 4.1, we suggested that *RWTS* has the potential to provide a *higher diversity* to a population than *TS* at a similarly high selection pressure, where diversity is defined as the number of different solutions in a population after the selection operation. This implies that *RWTS* has a higher

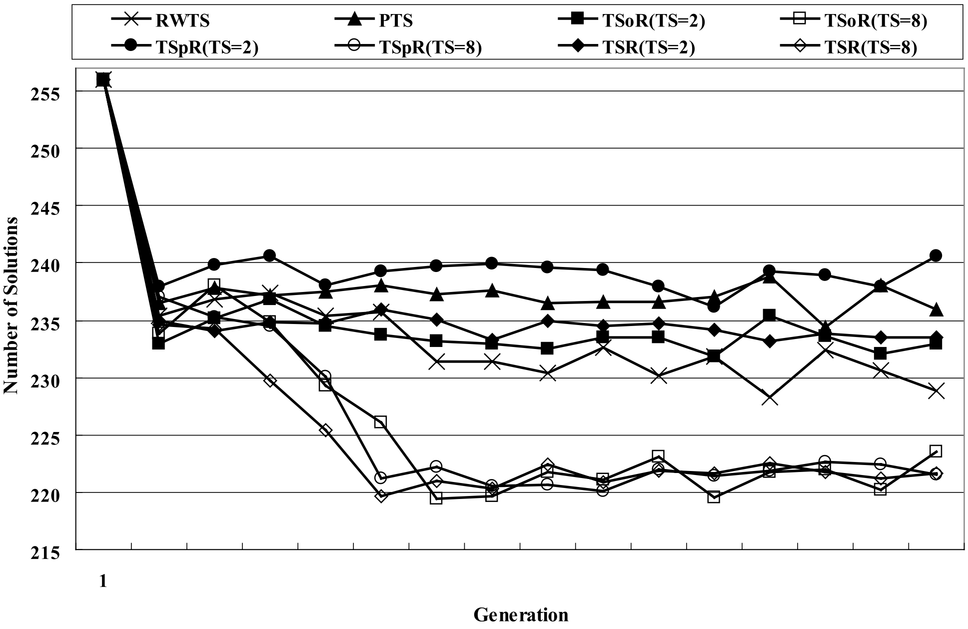


Fig. 5 Comparison of the diversity among selection strategies. Here, t_s indicates *tournament size*

possibility than *TS*; that solutions ranging from low fitness to high fitness can coexist in the same population. In this subsection, we performed computational experiments based on diversity to illustrate this fact. The experimental environment and test problems are the same as in Sect. 4.3 (i.e., the performance test), except for the initialization process of GAs. During initialization, the initial solutions group (i.e., the initial population) is identically generated for all tests in order to obtain a more objective comparison of the selection strategies.

Figure 5 presents the results of the diversity tests. In the figure, the subscript of *TS* denotes the tournament size t_s, and the high value in the vertical axis indicates high diversity. Each result in these selection strategies is the average of the results of 5 test problems, and each test problem result is the average of 30 independent trials. In a comparison of *RWTS* and *TS* with a tournament size of 2 ($t_s = 2$), *RWTS* shows the lowest diversity, but similar to the previous statistical analysis when $t_s = 8$, *RWTS* shows the highest diversity. This result indicates that *RWTS* can maintain a high population diversity with respect to a similar selection pressure as *TS*. Similar results were obtained from the tests with population sizes of 128, 512, and 1024.

The use of an excessively high pressure is not appropriate in the selection strategy because it could drive the selection to a situation of premature convergence. From these results, we can infer that *RWTS* does not have an excessively high selection pressure because there is only a small loss of diversity.

4.5 Discussion

From the above four types of empirical analysis, we could identify several interesting facts regarding the comparison of the characteristics of *RWTS* and *TS*, and also regarding the relationship between the characteristics of the selection strategies and the performance of GAs. The five points are summarized as follows:

(1) *RWTS* has a higher *selection pressure* than *TS* when $t_s = 2$.
(2) *RWTS* maintains more *diversity* than *TS* in conditions of *similar selection pressure*.
(3) *RWTS* has the highest *sampling accuracy* than *TS*, except for *TSpR* when $t_s = 2$.
(4) Comparatively high *election pressure* in the selection strategy provides good performance.
(5) High *sampling accuracy* in a selection strategy provides good performance.

Point 1 is obviously true, and is easily proved by comparing the selection probability distributions between *RWTS* and *TS* when $t_s = 2$. Point 2 can be inferred from the comparison of the selection probability distributions between *RWTS* and *TS* in conditions of similar selection pressure, and is confirmed by the diversity tests on a traveling salesman problem. Point 3 is determined from the stochastic tests. Point 4 could be deduced from the performance tests, especially, in the comparison of the performance of *TS* between $t_s = 2$ and $t_s = 8$, except for *TSpR*. The reason for the exception is that *TSpR* has a different sampling accuracy between $t_s = 2$ and $t_s = 8$. However, this point must be carefully considered because an excessively high pressure is not acceptable; keep in mind that this point is true only after being confirmed in the performance tests. Point 5 could be deduced from the performance tests, by comparing the performance of *TSpR* with that of other *TS*s.

From Points 1 and 2, we can view *RWTS* as being more balanced between exploration (i.e., concerned with diversity) and exploitation (i.e., concerned with selection pressure) than a conventional *TS* ($t_s = 2$). We suggest that this balance and the higher sampling accuracy of *RWTS* provides a better performance than *TS*, except for *TSpR*; *TSpR* shows a slightly better performance than *RWTS*, potentially due to its higher sampling accuracy than *RWTS*.

5 Conclusion

RWTS was successfully implemented and applied to various benchmark problems [5–8]. In this research, we have focused on statistical analysis of *RWTS* and *TS* from the viewpoint of the selection probabilities and stochastic properties in order to rationally explain what provides a good performance in the selection strategy. From four empirical analyses, it was demonstrated that a comparatively high selection pressure and high sampling accuracy, which are inherent in *RWTS*, can provide good performance in the selection strategy. Furthermore, it was shown that *RWTS* provides more diversity than *TS* under the condition of similar selection pressures. However, from the performance test, we found that the tournament selection with partial replacement gives a slightly better performance than *RWTS*. Then, based on the collective analysis of all empirical results, we suggest that this result might come from *sampling accuracy*.

Moreover, we have demonstrated how each characteristic of the selection strategy affects the performance. As a future work, we will investigate the correlation between the characteristics of the selection strategy and its performance.

Acknowledgements The authors would like to thank the referees for their valuable suggestions for improving the clarity and content of this paper. This research was financially supported by the UCN Project, the MIC 21C Frontier R&D Program in Korea, and the Center for Distributed Sensor Networks at GIST.

References

1. Holland JH (1975) Adaptation in natural and artificial system. University of Michigan Press, Ann Arbor
2. Baker JE (1985) Adaptive selection methods for genetic algorithms. In: Proceedings of the 1st international conference on genetic algorithms and their applications, pp 101–111
3. Michalewicz Z (1992) Genetic algorithms + data structures = evolution programs. Springer, Berlin
4. Goldberg DE, Klösener KH (1991) A comparative analysis of selection schemes used in genetic algorithms. In: Rawlins G (ed) Foundation of genetic algorithms. Kaufmann, San Mateo, pp 69–93
5. Soak S, Corne D, Ahn B (2004) A powerful new encoding for tree-based combinatorial optimisation problems. In: Lecture notes in computer science, vol 3242, pp 430–439
6. Soak S, Corne D, Ahn B (2004) A new encoding for the degree constrained minimum spanning tree problem. In: Lecture notes in computer science, vol 3213, pp 952–958
7. Soak S, Corne D, Ahn B (2006) The edge-window-decoder representation for tree-based problem. IEEE Trans Evol Comput 10(2):124–144
8. Lee S, Soak S, Mahalik NP, Ahn B, Jeon M (2006) Mathematical and empirical analysis of the real world tournament selection. In: Lecture notes in artificial intelligence, vol 4215, pp 130–137
9. Julstrom BA (1999) It's all the same to me: Revisiting rank-based probabilities and tournaments. In: Proceedings of the congress on evolutionary computation, pp 1501–1505
10. Bäck T (1996) Evolutionary algorithms in theory and practice. Oxford University Press, Oxford
11. Baker JE (1987) Reducting bias and inefficiency in the selection algorithm. In: Grefenstette JJ (ed). Proceedings of the second international conference on genetic algorithms. Kaufmann, San Mateo, pp 14–21
12. Beasley D, Bull DR, Martine RR (1993) An overview of genetic algorithms, part 1: fundamentals. University computing, the bulletin of the IUCC, vol 15, no 2, p 58
13. Rogers A, Bennett AP (1999) Genetic drift in genetic algorithm selection schemes. IEEE Trans Evol Comput 3(4):298–303
14. Schell T, Wegenkittl S (2001) Looking beyond selection probabilities: adaptation of the χ^2 measure for the performance analysis selection methods in GAs. Evol Comput 9(2):243–256
15. Blickle T, Thiele L (1995) A comparison of selection schemes used in genetic algorithms. Technical Report 11, Swiss Federal Institute of Technology, Zürich, Switzerland
16. Blickle T, Thiele L (1995) A mathematical analysis of tournament selection. In: Proceedings of the 6th international conference on genetic algorithms, pp 9–16
17. Bäck T (1994) Selective pressure in evolutionary algorithms: a characterization of selection mechanisms. In: Proceedings of 1st IEEE conference on evolutionary computation, pp 56–62
18. Papoulis A, Pillai SU (2002) Probability, random variables and stochastic processes, 4th edn. McGraw–Hill, New York, p 361
19. http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/index.html
20. Bean JC (1994) Genetics and random keys for sequencing and optimization. ORSA J Comput 6:154–160
21. Sywerda G (1989) Uniform crossover in genetic algorithms. In: Proceedings of the third international conference on genetic algorithms, George Mason University, USA, pp 2–9